Application of seismic attributes and neural network for sand probability prediction — A case study in the North Malay Basin

Jianyong Hou[1]*, Toshihiro Takahashi[1], Arata Katoh[1], Suwit Jaroonsitha[2], Khun Puvanat Chumsena[2] & Kazuo Nakayama[3]

[1]JGI, Inc., 1-5-21 Otsuka Bunkyo-ku, Tokyo, 112-0012, Japan
*Tel: +81-3-5978-8038, Fax: +81-3-5978-8050, E-mail address: hou@jgi.co.jp
[2]CPOC, PETRONAS Twin Towers, Kuala Lumpur City Centre, 50088 Kuala Lumpur, Malaysia
[3]JAPEX Co, Ltd., 1-7-12, Marunouchi Chiyoda-ku, Tokyo, 100-0005, Japan

Abstract— The application of seismic data has greatly increased our capability to correctly understand the distributions of reservoir layers and geological structures of interest. Seismic attributes analysis is a popular and important method to predict the distributions of reservoir properties such as lithology, porosity and thickness. In the North Malay Basin, some sandstone layers related to reservoirs are found between A and B horizons. Especially, the gas reservoirs have been confirmed in the sandstone layers around Z horizon developed between A and B. Although 3D seismic survey has been executed, the distributions of the sandstone layers, e.g. thickness, porosity, still remain unclear across the area because there are only few wells drilled. To understand the distributions of sandstone layers and lithologic change across the whole 3D seismic survey area, we utilized Artificial Neural Network (ANN) method supported by Geology Driven Integration (GDI) technique. This process established the relationship between sandstone lithology and some seismic attributes. We selected Z horizon as calculating datum plane to predict the sandstone distribution and lithologic change within the range from 150 ms above the horizon to 120 ms below it, that is from A horizon to B horizon.

Keywords: Artificial Neural Network, seismic attributes, rock properties, pseudo well, relationship

INTRODUCTION

Well logs and seismic records can be thought as responses to change of material properties such as porosity, lithology, fluid content and bed thickness in subsurface rocks. In general, the most effective and reliable method to obtain the rock properties is well log analysis. Because well log is 1D data, however, the analysis can only obtain the properties of underground rock or stratigraphy at well location and its neighborhood. Therefore, when only a few well log data are available (it is usual in the early stage of exploration) or when the wells concentrate in a part of study area, it is usually difficult to understand the distributions of rock properties across the whole study area. On the other hand, each sample in a seismic trace has given amplitude, frequency, phase, etc. and seismic records, 2D or 3D, contain usually very abundant information of lateral and vertical change of rock properties. Therefore, seismic records are closely related to well logs. If a mathematical relationship can been established between some seismic attributes from seismic records and rock properties from well logs, we can apply the relationship to predict the distributions of rock properties away from well control in 2D or 3D seismic survey area. The Artificial Neural Network (ANN) method supported by the Geology Driven Integration (GDI) technique provides such a tool for us.

GDI CONCEPT AND THEORY

The GDI is a technique for predicting distribution and change of rock properties, *e.g.* thickness, porosity, etc. directly from the seismic attributes (Nakayama & Hou, 2002, Nakayama *et al.*, 2003).

In GDI, ANN method is utilized to establish the mathematical relationship. A problem, however, is to collect a large number of well data to establish a statistically significant relationship. Only a few available well data in the early stage of exploration had prevented us from establishing the relationship. To compensate such a condition under a few wells, a large number of pseudo wells can be generated by a method of Monte Carlo Simulation in GDI (de Groot *et al*., 1996).

The Monte Carlo Simulation is a procedure that involves sampling based on probability to approximate the solution of mathematical or physical problems in a statistical way. There are two major advantages in this method superior to conventional ways. The first advantage is that we can steer the algorithm with rules based on geological reasoning. The second is that we can include hard constraints for each of the stochastic variables. In GDI, the Monte Carlo method is used to generate pseudo wells based on the stratigraphic-geological model built from factual well data, geological information and geological knowledge. These pseudo wells can be considered as an approach to the geological realities of study area. Only purpose that we generate these pseudo wells is to compensate the lack of lithologic samples from factual wells. Therefore, these pseudo wells generated by Monte Carlo Simulation have various stratigraphic frameworks and well log responses, but have no information on spatial locations.

ANN, or sometimes referred to as connection model, has emerged in the last decade as a promising computing technique and has been applied successfully in a variety of scientific and technological fields. It simulates the cognitive processes of the human brain and is suited for solving difficult problems, such as character recognition. In our study, ANN is used to establish the relationship between rock properties and seismic attributes. The ANN as non-linear dynamic system can find the relationship thorough training it with some known samples from well logs and corresponding seismic traces.

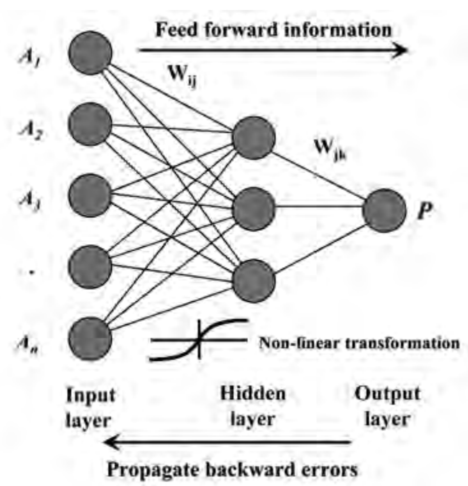

Figure 1 : Architecture of Multi-layers perceptron.

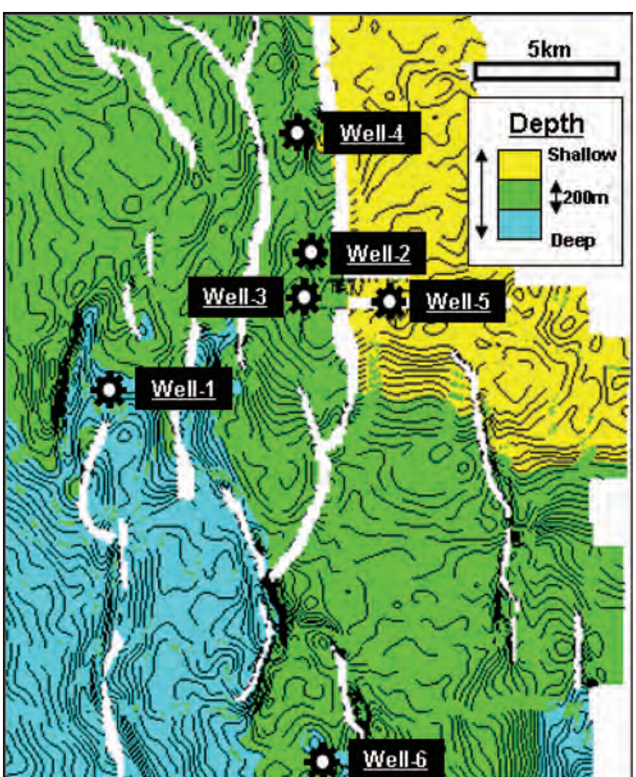

Figure 2: Depth structure map of top Z-sand Reservoir in the study area.

The most general ANN model is the multi-layers perceptron with the learning algorithm of back-propagation. It is constituted always of more than three layers, or an input layer which contains several input nodes, an output layer which contains one or several output nodes, and at least one hidden layer which contains several nodes (Figure 1). Nodes in the hidden and output layers are sometimes referred to as processing units. Nodes of adjacent layers are interconnected by weighting vectors that are initially randomized, but there are no connections between the nodes belonging to the same layer. The training process of back-propagation ANN involves sending the input values forward through network, and then computing the difference between the calculated output and the corresponding desired output from factual and pseudo wells. The error information is propagated backwards through the back-propagation. The mathematical expression of back-propagation ANN is represented by the following equations:

$$x = \sum^{n} W_i \times A_i$$

$$P = \frac{2}{1+e^{-x}} - 1$$

where *W* is the weighting vector, *A* is the input vector of neural network (seismic attribute), *x* is a weighting function, *P* is output (rock property) which is calculated by a sigmoid exponential function, and n is number of input vectors. The training algorithm attempts to minimize the error between the computed output and the desired output values by automatically adjusting the weighting vectors of connections (Rumelhart *et al*., 1986). When the errors fail to further reduce, the training is finished and the weighting vectors that yield the minimum error are saved, and the functions are determined.

GEOLOGICAL BACKGROUND OF STUDY AREA

Study area is geologically located in the North Malay Basin in the offshore southern part of the Gulf of Thailand. The North Malay Basin is a Neogene intracratonic rift/sag basin, elongated in shape, oriented in the northwest-southeast direction and consists of provinces of margin platform, margin ramps and central basinal area or depocenter. Geochemical studies indicate a high source rock potential for gas and also for substantial amount of oil. Hydrocarbons within the Malay Basin are found in all of the above provinces. Generally, the North Malay Basin contains thick sediments in excess of 8 km deposited since the Oligocene. Tecto-stratigraphically, the basin is divisible into a Pre-, Syn-, and Post-rift Megasequence. The Megasequence is further divided into Upper and Lower Post-Rift Megasequence and in some cases in the north divisible into Regional Sagging Megasequence. The structural styles in the study arae are controlled by the movement of normal faults, both pure extensional and transpressional tectonics. The extensional tectonics during Oligocene and Early Miocene caused tilted fault blocks, horst and graben features and anticlinal features against faults (Figure 2). The trap style is mainly fault dependent while widespread marine shales provide a good regional top seal. The intraformational shale in the prospective sequence also provides a very good top seal for individual reservoir package. In the study area, some sandstone layers with reservoirs property are found between A and B horizons, and main reservoir is Z-sand between them (Figure 3).

APPLICATION OF GDI

Usually, the lithology information can be obtained from log data. In our study area, however, it is difficult to get it because there are six available wells and only three among them can be used in the study. Therefore, for understanding the distributions of the sandstone layers between A and B horizons, the GDI technique is applied in the study area.

a. Pseudo wells

To predict the distribution of the sandstone using the seismic attributes, it is necessary to establish the mathematical correlation, or relationship, between seismic attributes and lithologic property. Obviously, it is difficult to establish the statistical relationship only by limited three factual well data. To obtain enough lithologic samples, we need more well data. For this reason, a 1D simulation called the Monte Carlo Simulation is introduced to generate pseudo wells to compensate the lack of factual well data (de Groot *et al.,* 1996).

A geological model is first established by integrating various real geological data including well logs, geological information and knowledge of the study area (Figure 4) before generating pseudo wells. In the geological model, the stratigraphic order, lithologic units and their properties in several formations are defined and are attached with various log responses, and various lithologies are distinguished from relative GR value. For sampling the problem, we only define two kinds of lithologic units in every formation unit, or sandstone and shale. Moreover, the sand and shale units are given a lothologic code respectively to distinguish and classify them with seismic attributes. That is, sand is defined as one (1) and shale as zero (0).

To obtain enough well realizations, 200 pseudo wells with varying stratigraphic frameworks and well log responses are generated (Figure 5 left). In this simulation, the parameters of variables such as thickness of stratigraphic units, porosity and various log responses of lithologic units are changed simultaneously and stochastically, based on the statistics of the factual well data and given geological constraints. Such pseudo wells have various log responses and are considered as similar to geological realities of the study area, and therefore, can be used as known samples of litholgic properties, but they have not location information. Only purpose to generate pseudo wells is to compensate the lack of factual wells. Therefore, the location information is not necessary in the simulation.

The synthetic seismograms corresponding to the pseudo wells can be constructed by convolving the reflection coefficient series with a zero-phase wavelet extracted from the real 3D seismic records (Figure 5 right). The pseudo wells are an approach to lithologic properties of the study area, while the synthetic seismograms represent the feature of seismic attributes.

b. Establishment of the mathematical relationship, utilizing ANN

As stated before, both well logs and seismic attributes may be considered as a response to various rock properties. We can obtain the rock properties, such as lithology, thickness, porosity, fluid-bearing character, etc. from factual and/or pseudo well logs, and seismic attributes from seismic records or from synthetic traces. Since the relationship is not necessarily linear, the ANN method that recognizes non-linear correlation is used in GDI analysis.

We train the ANN by feeding some known samples of lithology and seismic attributes extracted from pseudo wells and synthetic traces respectively into the ANN. By the training, the ANN can find a non-linear correlation

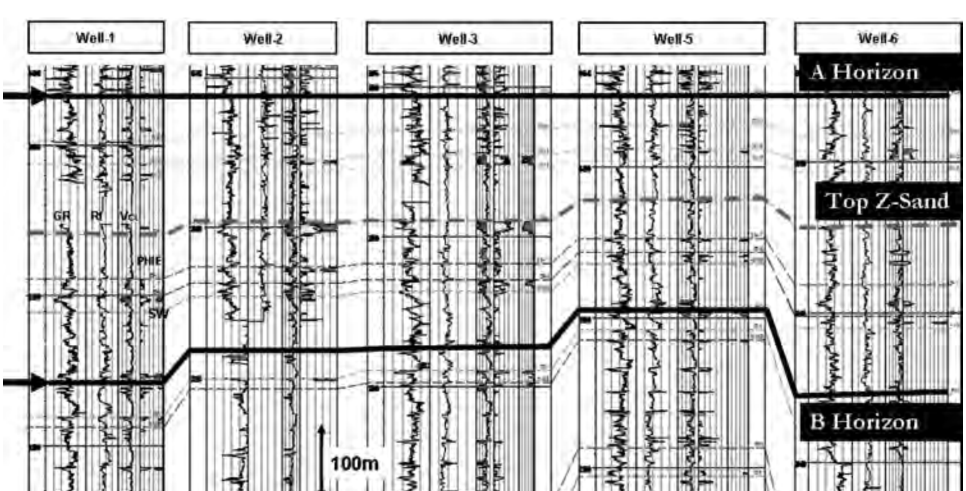

Figure 3: Well correlation among wells in the study area. Main sand reservoirs are found around Z horizon.

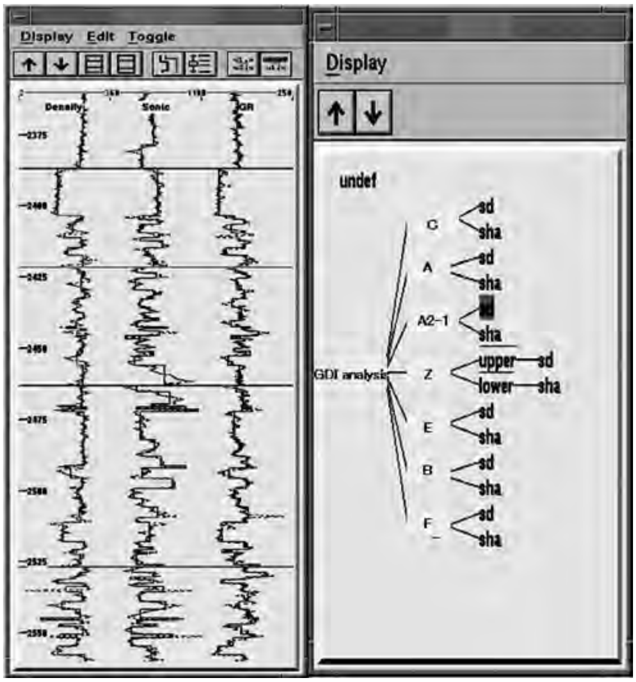

Figure 4: Geological-stratigraphic model established from real geological information including well logs, stratigraphic sequence, lithologic units and geological knowledge.

between lithology and seismic attributes. In this study, used seismic attributes, called "Sample", is a time series of amplitude values along a given seismic trace. In other words, it returns the amplitudes at the specified sampling positions within a specified time windows (Figure 6). When "Sample" is extracted, the extracting sample rate, the time window length at every sampling and total sampling time window length are necessarily specified. In this case, the sampling rate is 2ms and the time window length at every sampling is defined from -20 ms to 20 ms referring to sampling position, while the total sampling time window length is from -150 ms to 120 ms referring to Z horizon (Figure 6), which corresponds approximately from A horizon to B horizon (Figure 3).

There are several ANN models, but the Multi-layers perceptron model with the learning algorithm of back-propagation is used to establish the relationship between the lithology codes and seismic attributes (Figure 7). The advantage of the model is that it can feed back the error information to input layer by comparing the calculated values (lithology property) by ANN with "actual" values extracted from wells (including factual and/or pseudo wells), and then adjust automatically the weighting vectors interconnecting

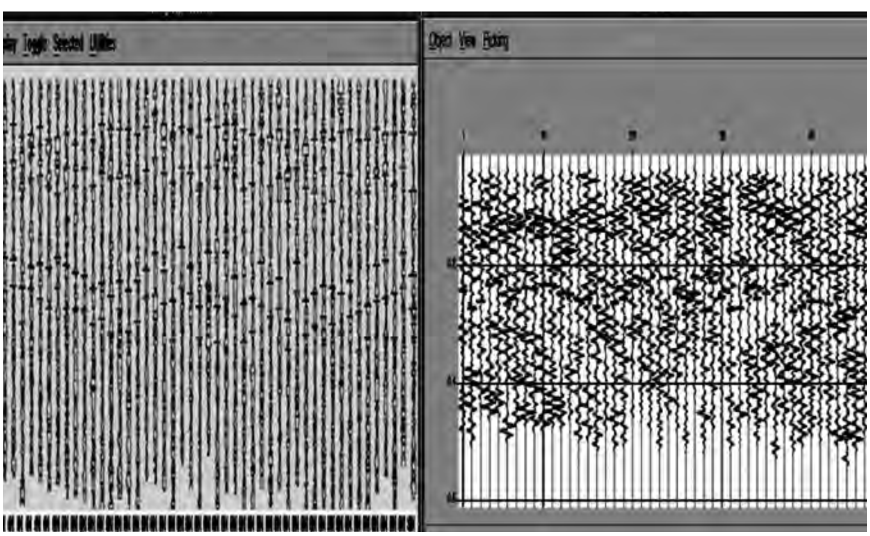

Figure 5: Part of pseudo wells (left) generated by Monte Carlo simulation and corresponding synthetic traces (right), based on the geological-stratigraphic model.

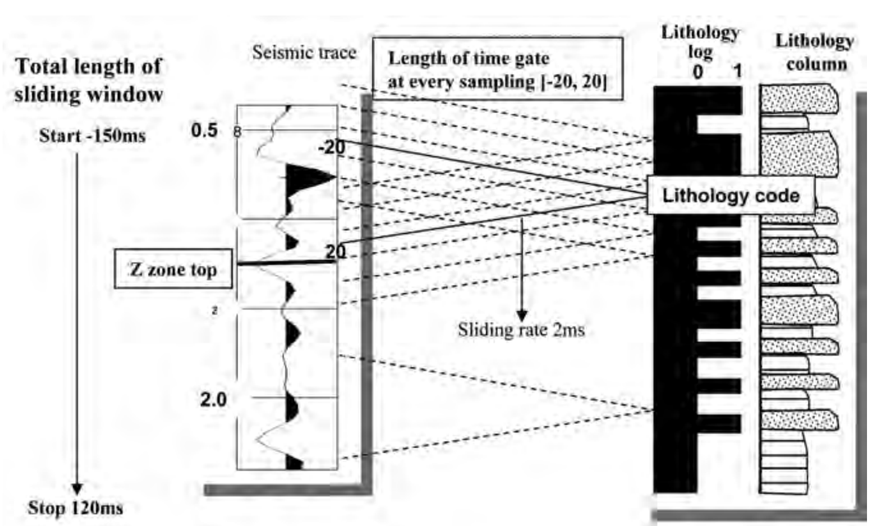

Figure 6: Lithology extraction from pseudo wells and seismic attributes (Sample or Amplitude) from synthetic traces as known samples.

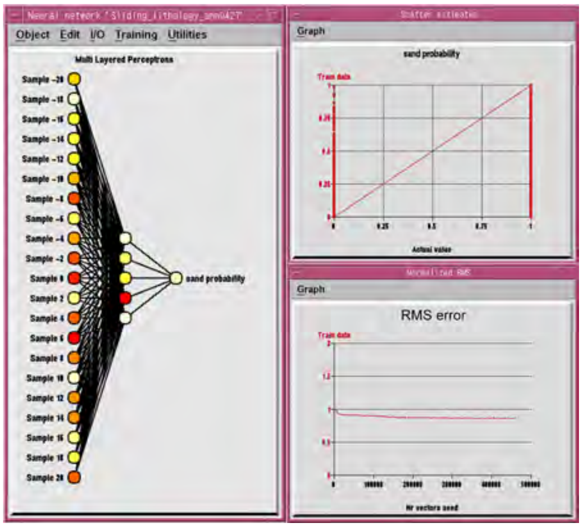

Figure 7: The trained ANN (left) and its performances (right) for establishing the relationship between the lithologic code and seismic attributes. Nodes are displayed in color scale relative to their contribution to the output, i.e., from red (highest) , orange, yellow to white (lowest).

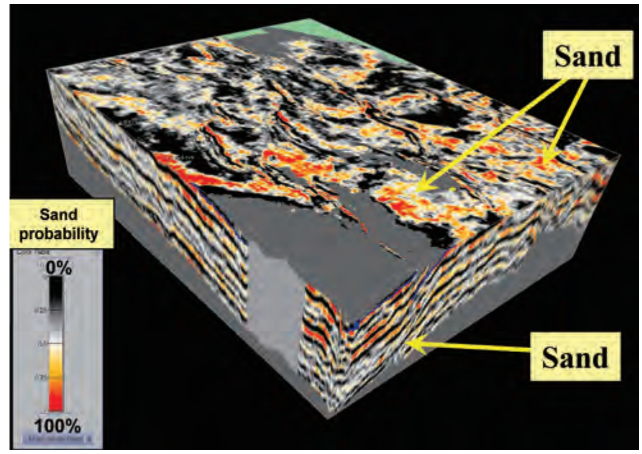

Figure 8: A lithology cube created by applying the ANN to 3D seismic volume. It shows the distribution of sand probability from A horizon to B horizon within the seismic volume.

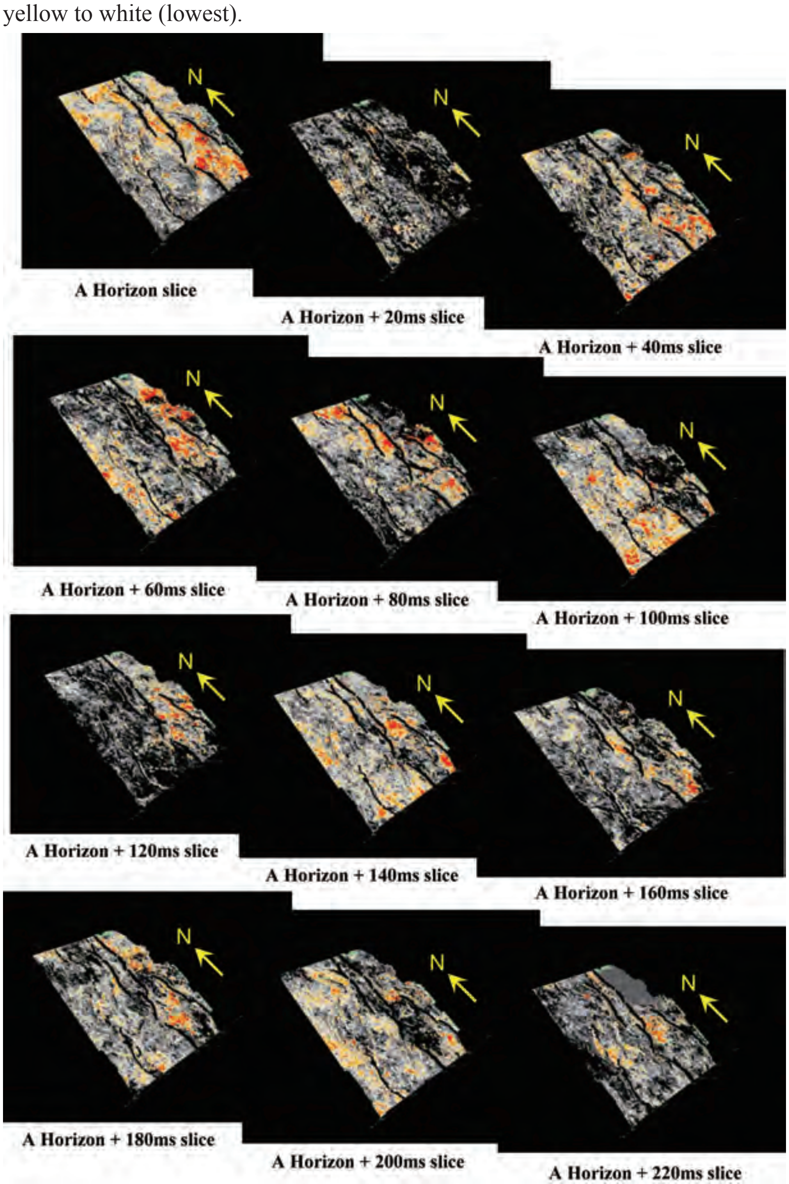

Figure 9: The distribution of sand probability on the horizon slices with different time.

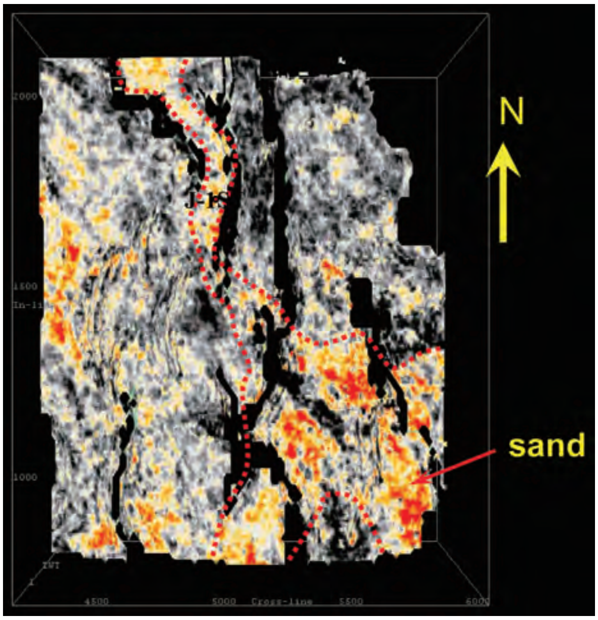

Figure 10: Sand distribution on the Z horizon slice.

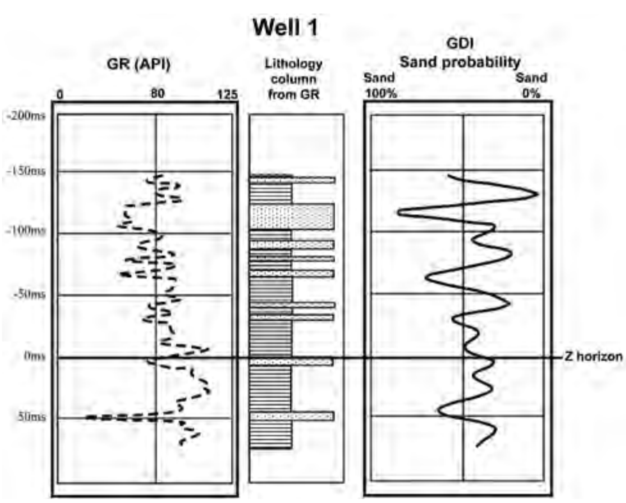

Figure 11: Comparison of predicted lithology by GDI with GR from factual well 1 at well locations.(Solid line: sand probability, Dotted line: GR).

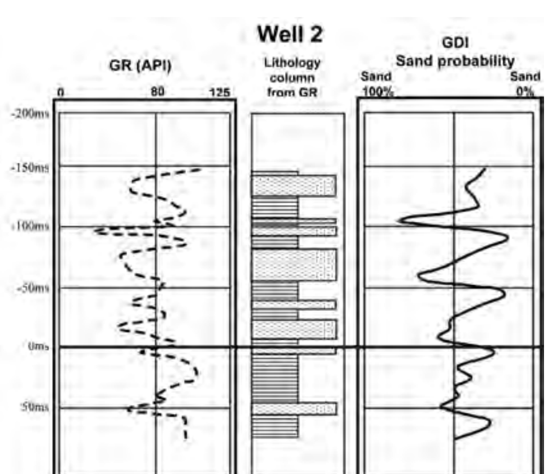

Figure 12: Comparison of predicted lithology by GDI with GR from factual well 2 at well locations.(Solid line: sand probability, Dotted line: GR).

adjacent layer so that the RMS errors can be decreased to a acceptable or reasonable level. According to the experiences, a RMS error below 0.8 is considered generally to be reasonable. The right lower figure in the Figure 7 shows the decay of the calculative normalized RMS error during training. The final error is about 0.8. Although it is a little high, it is still thought in a reasonable range.

c. Prediction of distribution of sand probability

The trained ANN is applied to a 3D seismic volume in the North Malay Basin to predict the distribution of sand probability within the interval about from A horizon to B horizon.

The Figure 8 is a lithology cube showing the distribution of sand probability in the specified interval of the 3D seismic survey. It is created by applying the trained ANN above to the 3D seismic volume of the North Malay Basin. The Figure 9 shows the distribution of sand probability on several horizon slices with different depth, which is generated parallel to the A horizon and each 20ms interval. From these illustrations on the Figure 9, we can find there are little sand probability on some horizon slices, *e.g.* A horizon+20 ms, A horizon+120 ms and A horizon+160 ms, etc. This feature suggests some predominant zones of shale or argillaceous rock around these horizon slices. Moreover, on the Z horizon slice (Figure 10), the distribution of sand probability shows a channel and fan in shape. This feature may indicate the deposit environment of sandstone. On the lithology sections along Inline or Crossline (Figure 8), the most of the high sand probability zones (red color) are observed in discontinuous distributions. This may indicate that sandstone develops in discontinuous sand body or lens.

CONCLUSIONS

The Geology Driven Integration Tool including the artificial neural network and Monte Carlo simulation techniques is applied to the 3D seismic survey of the North Malay Basin to predict the distribution of sand probability within the main reservoir horizons. The predicted result is checked with GR from factual wells (Figures 11a and 11b). The check shows that the predicted result is consistent with log features on the whole although there is error to some extent.

Using the seismic attributes and ANN technique is an effective method in the prediction of distributions of rock properties, especially when factual well data are few or too concentrated so that it is difficult to understand the distributions by conventional methods.

ACKNOWLEDGEMENTS

We thank Carigali-PTTEP Operating Company Sdn. Bhd (CPOC) and Malaysia-Thailand Joint Authority (MTJA) for permission to publish this paper, which is based on the data and information of seismic and geology from a field situated in the North Malay Basin.

REFERENCES

De Groot, P.F.M.,. Bril, A.H., Floris, F.J.T. and Campbell, A.E. 1996, Monte Carlo simulation of wells, Geophysics, 61(3), 631-638.

Nakayama, K. and Hou, J., 2002, Prediction of Reservoir Properties by Monte Carlo Simulation and Artificial Neural Network in the Exploration Stage. Soft Computing for Reservoir Characterization and Modeling. Edited by P.Wong, F.Aminzadeh and M.Nikravesh, pp. 15-33.

Nakayama, K., Hou, J., Ohkawa, S. and Mitsuda, S. 2003, Application of Geology Driven Integration Technique in the Prediction of Reservoir Distribution: An Example from the Japan Sea. The paper collection of the 9th Formation Evaluation Symposium of Japan, SPWLA.

Rumelhart, D.E., Hinton, G.E. & Williams, R.J., 1986. Learning internal representation by error propagation. In: Rumelhart, D. E. & McClelland, J. L. (eds.) Parallel Distributed Processing. Cambridge, MIT Press, pp. 318-362.

Manuscript received 28 May 2008
Revised manuscript received 25 November 2008